

Algebraic theory for higher-order methods in computational quantum mechanics

Pranav Singh*

November 26, 2015

Keywords: Free Lie algebras, derivations, Jordan products, Lie group methods, exponential splittings, Zassenhaus splitting, Magnus expansion, semiclassical Schrödinger equations

Abstract

We present the algebraic foundations of the *symmetric Zassenhaus algorithm* and some of its variants. These algorithms have proven effective in devising higher-order methods for solving the time-dependent Schrödinger equation in the semiclassical regime. We find that the favourable properties of these methods derive directly from the structural properties of a \mathbb{Z}_2 -graded Lie algebra. Commutators in this Lie algebra can be simplified explicitly, leading to commutator-free methods. Their other structural properties are crucial in proving unitarity, stability, convergence, error bounds and quadratic costs of Zassenhaus based methods. These algebraic structures have also found applications in Magnus expansion based methods for time-varying potentials where they allow significantly milder constraints for convergence and lead to highly effective schemes. The algebraic foundations laid out in this work pave the way for extending higher-order Zassenhaus and Magnus schemes to other equations of quantum mechanics.

1 Introduction

Recently devised symmetric Zassenhaus splittings (Bader, Iserles, Kropielnicka & Singh 2014) and Magnus–Zassenhaus methods (Iserles, Kropielnicka & Singh 2015, Bader, Iserles, Kropielnicka & Singh 2015) for computationally solving semiclassical time-dependent Schrödinger equations (TDSEs) have proven highly effective in achieving arbitrarily high orders of accuracy with costs growing quadratically in the desired order. This is in contrast to Yoshida based methods (Yoshida 1990), whose costs grow exponentially with order. These new schemes are commutator-free and allow much larger time steps than otherwise thought possible for convergence of Magnus and Zassenhaus expansions.

Underlying the effectiveness of these new methods are some algebraic structures which form the crux of this work. In a break from tradition in numerical solutions of partial differential equations, the methods of (Bader et al. 2014) are devised by working directly in the *free Lie algebra* of undiscretised operators ∂_x^2 and $V(x)$. The quadratic costs of symmetric Zassenhaus methods are achievable due to the property of *height reduction* in these Lie algebras. This property entirely fails to materialise when working with matrices corresponding to

*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Rd, Cambridge CB3 0WA, UK. E-mail: singh.pranav@gmail.com, Tel: +44 (0)1223 337892, Fax: +44 (0)1223 765900.

the discretisation of ∂_x^2 and $V(x)$. Hence the structure of these free Lie algebras is crucial to the effectiveness of Zassenhaus schemes.

The characterisation of these algebras forms a major motivation of this work. It is this characterisation that will allow us to formally prove the quadratic growth in costs. We will also find that these algebras possess a \mathbb{Z}_2 -graded structure that, along with the symmetry of its constituents, proves crucial for the numerical stability as well as the unitary evolution of the wave-packet. Instead of characterising the free Lie algebra,

$$\text{FLA}(\{V(x), \partial_x^2\}),$$

however, we will attempt the characterisation of a bigger free Lie algebra,

$$\text{FLA}(\mathcal{S} \cup \mathcal{P}(\partial_x)),$$

generated by all functions in the function space \mathcal{S} and all polynomials of the differential operator ∂_x .

These algebraic characterisations will be introduced in a more abstract form—in the context of an associative algebra \mathcal{A} with a commutative subalgebra \mathcal{C} and its Lie idealiser \mathcal{I} . The algebraic structures of the Zassenhaus schemes for the TDSE are seen to be special cases of the Lie algebras introduced here. To be more concrete, the associative algebra \mathcal{A} in the case of the TDSE will consist of linear operators, the commutative subalgebra will be formed by multiplicative operators (whose action is multiplication by V , for instance), and ∂_x will be seen to be an element of the Lie idealiser \mathcal{I} . The properties of these Zassenhaus schemes are traced to the structural properties of these algebras.

This algebraic formulation will prove of crucial importance in cutting the Zassenhaus algorithms free of their semiclassical TDSE origins, allowing their use in devising higher-order computational methods for other equations of quantum mechanics and possibly beyond.

In Section 2 we introduce the abstract context in which our algebra \mathfrak{F} is defined. In Section 3, we show that \mathfrak{F} is an associative algebra and study its structure. While it follows immediately that \mathfrak{F} is also a Lie algebra, in Section 4 we find out that it possesses a very interesting structure which has significant ramifications for numerical methods discussed in Section 5.

The tables of coefficients in Appendix A should aid direct applications of the results of this paper. The lengthier proofs of Lemma 2 and Lemma 3 from Section 3 have been confined to Appendices B and C, respectively, for ease of narration.

2 Some notations and definitions

Consider a commutative algebra \mathcal{C} which is a subalgebra of the unital associative algebra $(\mathcal{A}, \cdot, +)$ over the field of rational numbers \mathbb{Q} . The commutator on the associative product,

$$[a, b] = a \cdot b - b \cdot a,$$

acts as the canonical Lie product while the anticommutator,

$$a \bullet b = \frac{1}{2}(a \cdot b + b \cdot a),$$

acts as a Jordan product. \mathcal{A} along with the Lie (Jordan) product forms a Lie (Jordan) algebra which we can identify with \mathcal{A} again. The Lie idealiser of \mathcal{C} in \mathcal{A} ,

$$\mathcal{I} = \{d \in \mathcal{A} : [d, \mathcal{C}] \subseteq \mathcal{C}\}.$$

is the largest subalgebra of \mathcal{A} in which \mathcal{C} is a Lie ideal.

In this work we typically use a, b, c for elements of the associative algebra \mathcal{A} , x, y, z for elements of the commutative subalgebra \mathcal{C} and reserve d for elements of the Lie idealiser \mathcal{I} , except where it is clear from context. For elements x, y of the commutative algebra \mathcal{C} , we write xy to denote $x \cdot y$, dropping the explicit use of the multiplication operator. The letters k, l, m, n, p, r, s, i are used for non-negative integers.

Consider the adjoint map, where the action of ad_a for $a \in \mathcal{A}$ is described by

$$\text{ad}_a(b) = [a, b].$$

ad_a is always a derivation on \mathcal{A} and is called an inner derivation. However, it need not be a derivation on \mathcal{C} in general. For $d \in \mathcal{I}$, ad_d is also a derivation on \mathcal{C} ,

$$\text{ad}_d \in \text{Der}(\mathcal{C}), \quad d \in \mathcal{I}.$$

Definition 1 For $x \in \mathcal{C}, d \in \mathcal{I}$ and a non-negative integer k we introduce the notation

$$\langle x \rangle_k^d = x \bullet d^k = \frac{1}{2}(x \cdot d^k + d^k \cdot x),$$

where the power d^k is defined as usual, $d^{k+1} = d \cdot d^k$ and $d^0 = 1_{\mathcal{A}}$.

Definition 2 For any $d \in \mathcal{I}$ and non-negative integer k , we define the linear space

$$\mathfrak{F}_k^d = \left\{ \langle x \rangle_k^d : x \in \mathcal{C} \right\},$$

and denote the direct sum of these linear spaces by

$$\mathfrak{F}^d = \bigoplus_{k \in \mathbb{Z}_0^+} \mathfrak{F}_k^d.$$

In this work we will only use a single non-trivial element of \mathcal{I} which is not in \mathcal{C} , $d \in \mathcal{I} \setminus \mathcal{C}$, and can drop the superscript d , writing $\langle x \rangle_k$, \mathfrak{F}_k and \mathfrak{F} in place of $\langle x \rangle_k^d$, \mathfrak{F}_k^d and \mathfrak{F}^d , respectively. To simplify notation further, we denote $\text{ad}_d^i(x)$ by $D^i x$.

Since D is a derivation on \mathcal{A} (and \mathcal{C}), $D(a \cdot b) = Da \cdot b + a \cdot Db$, and it distributes binomially on \mathcal{A} (and \mathcal{C}),

$$D^k(a \cdot b) = \sum_{i=0}^k \binom{k}{i} D^i a \cdot D^{k-i} b. \quad (2.1)$$

Starting from $d \cdot x = Dx + x \cdot d$, a simple inductive procedure leads to a similar binomial identity,

$$d^k \cdot x = \sum_{i=0}^k \binom{k}{i} D^i x \cdot d^{k-i}. \quad (2.2)$$

3 \mathfrak{F} as an associative algebra

Lemma 1 The linear space \mathfrak{F} is an associative algebra with

$$\langle x \rangle_k \cdot \langle y \rangle_l = \sum_{n=0}^{k+l} \langle z_n \rangle_{k+l-n}, \quad (3.1)$$

where the terms

$$z_n = \sum_{i=0}^n \pi_{n,i}^{k,l} D^i x D^{n-i} y \quad (3.2)$$

are in \mathcal{C} and $\pi_{n,i}^{k,l} \in \mathbb{Q}$.

Proof A proof for Lemma 1 follows by expanding (3.1) using the binomial identity (2.2) and comparing powers of d . At first, we expand the left hand side,

$$\begin{aligned} 4 \langle x \rangle_k \cdot \langle y \rangle_l &= (x \cdot d^k + d^k \cdot x) \cdot 2 \langle y \rangle_l \\ &= x \cdot (d^k \cdot y) \cdot d^l + x \cdot (d^{k+l} \cdot y) + (d^k \cdot x) \cdot 2 \langle y \rangle_l \\ &= x \cdot \left(\sum_{i=0}^k \binom{k}{i} D^i y \cdot d^{k-i} \right) \cdot d^l + x \cdot \sum_{i=0}^{k+l} \binom{k+l}{i} D^i y \cdot d^{k+l-i} \\ &\quad + \left(\sum_{i=0}^k \binom{k}{i} D^i x \cdot d^{k-i} \right) \cdot (y \cdot d^l + d^l \cdot y) \\ &= \sum_{i=0}^{k+l} \left[\binom{k}{i} + \binom{k+l}{i} \right] (x D^i y) \cdot d^{k+l-i} \\ &\quad + \sum_{i=0}^k \sum_{j=0}^{k+l-i} \binom{k}{i} \left[\binom{k-i}{j} + \binom{k+l-i}{j} \right] (D^i x D^j y) \cdot d^{k+l-i-j}. \end{aligned}$$

The right hand side of (3.1) is expanded in similar fashion,

$$4 \sum_{n=0}^{k+l} \langle z_n \rangle_{k+l-n} = 2 \sum_{n=0}^{k+l} z_n \cdot d^{k+l-n} + 2 \sum_{n=0}^{k+l} \sum_{i=0}^{k+l-n} \binom{k+l-n}{i} D^i z_n \cdot d^{k+l-n-i}.$$

A sequence of terms z_n for which all terms accompanying d^{k+l-a} , $a \in \{0, \dots, k+l\}$, match on both sides will certainly satisfy (3.1) although it need not be the unique solution. This leads to the relation

$$\begin{aligned} 2z_a + 2 \sum_{n=0}^{k+l} \binom{k+l-n}{a-n} D^i z_n &= \left[\binom{k}{a} + \binom{k+l}{a} \right] (x D^a y) \\ &\quad + \sum_{i=0}^a \binom{k}{a-i} \left[\binom{k-i}{a-i} + \binom{k+l-i}{a-i} \right] (D^i x D^{a-i} y), \end{aligned} \quad (3.3)$$

which can easily be fashioned into a recursive procedure for solving z_a , starting from $z_0 = xy$ (found by substituting $a = 0$).

The second part of the lemma states that z_n of the form (3.2) satisfy (3.1) for some $\pi_{n,i}^{k,l} \in \mathbb{Q}$. After substituting this form in (3.3), the left side is comprised of

$$2 \sum_{i=0}^a \pi_{a,i}^{k,l} D^i x D^{a-i} y \quad (3.4)$$

and

$$\begin{aligned} 2 \sum_{n=0}^{k+l} \binom{k+l-n}{a-n} D^{a-n} \left(\sum_{i=0}^n \pi_{n,i}^{k,l} D^i x D^{n-i} y \right) \\ = 2 \sum_{n=0}^{k+l} \binom{k+l-n}{a-n} \sum_{i=0}^n \pi_{n,i}^{k,l} \sum_{j=0}^{a-n} \binom{a-n}{j} D^{i+j} x D^{a-i-j} y, \end{aligned} \quad (3.5)$$

where the inner term $D^{a-n} \cdot (D^i x D^{n-i} y)$ has been expanded using (2.1).

We now equate terms accompanying $D^p x D^{a-p} y$ in (3.3), noting that this would give a solution which need not be unique. We arrive at equations of the form

$$R_{a,p}^{k,l} = L_{a,p}^{k,l}, \quad p \in \{0, \dots, a\}, \quad a \in \{0, \dots, k+l\},$$

where

$$R_{a,p}^{k,l} = 2\pi_{a,p}^{k,l} + 2 \sum_{n=0}^a \binom{k+l-n}{a-n} \sum_{i=0}^n \pi_{n,i}^{k,l} \binom{a-n}{p-i}, \quad (3.6)$$

$$\begin{aligned} L_{a,p}^{k,l} &= \delta_{p,0} \left(\binom{k}{a} + \binom{k+l}{a} \right) \\ &\quad + \binom{k}{p} \left(\binom{k-p}{a-p} + \binom{k+l-p}{a-p} \right), \end{aligned} \quad (3.7)$$

and $\delta_{i,j}$ is the Kronecker delta function. The fact that a recursive procedure for finding π can be designed,

$$\pi_{a,p}^{k,l} = \frac{1}{4} \left(L_{a,p}^{k,l} - 2 \sum_{n=0}^{a-1} \binom{k+l-n}{a-n} \sum_{i=0}^n \pi_{n,i}^{k,l} \binom{a-n}{p-i} \right), \quad (3.8)$$

starting from $\pi_{0,0}^{k,l} = 1$, is the proof that such $\pi_{n,i}^{k,l} \in \mathbb{Q}$ exist, whereby some $z_n \in \mathcal{C}$ of the form (3.2) satisfy (3.1). Hence \mathfrak{F} is an associative algebra with the prescribed properties (3.1) and (3.2). \square

3.1 Explicit form of the coefficients

The recursive procedure (3.8) suffices for the proof of Lemma 1 and for expanding products in the associative algebra \mathfrak{F} using a symbolic computation algorithm. Coefficients derived using this procedure for a few combinations of k and l are listed in Table 1 in Appendix A.

Nevertheless, an explicit form for these coefficients remains highly desirable for a variety of reasons. Firstly, this could give us an explicit form for some analytic functions on \mathfrak{F} (or some subspace of \mathfrak{F}), such as the exponential map. Secondly, an analysis of growth of these coefficients could allow us to deduce estimates concerning the convergence of some of these functions. Lastly, the observations which lead us to the structural properties of the Lie algebra in Section 4—the main result of this paper—are not immediately evident in the recursive form.

We state here some results concerning the form of the coefficients, confining the proofs to appendix B and appendix C for ease of narration.

Lemma 2 *The explicit form of the coefficients is given by*

$$\pi_{n,i}^{k,l} = \frac{1}{2} \sum_{s=0}^{k+l} \sum_{j=0}^{k+l} A_{(n,i),(s,j)}^{k+l} L_{s,j}^{k,l}, \quad (3.9)$$

where

$$A_{(n,i),(s,j)}^q = \begin{cases} \delta_{n,s} \delta_{i,j} - \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j} & n \geq s, i \geq j, \\ 0 & \text{otherwise,} \end{cases} \quad (3.10)$$

and P_r are defined in terms of the Bernoulli numbers B_r ,

$$P_r = (-1)^r (2^r - 1) B_r.$$

Some properties of the coefficients $\pi_{n,i}^{k,l}$ become evident once we study their generating functions. In this context we make an exception in notation, using u, w, x and y as variables in which the formal series of the generating function is specified.

Lemma 3 *The generating function,*

$$h(u, w, y, x) = \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \pi_{n,i}^{k,l}, \quad (3.11)$$

for the coefficients $\pi_{n,i}^{k,l}$ appearing in (3.1) is

$$h(u, w, y, x) = \frac{\exp((wy - uxy)/2)}{1 - (w + u)} \frac{\cosh(uy/2) \cosh(wxy/2)}{\cosh(y(u + w)(1 + x)/2)}. \quad (3.12)$$

Lemma 4 *The coefficients possess the symmetry,*

$$\pi_{n,i}^{k,l} = (-1)^n \pi_{n,n-i}^{l,k}.$$

Proof We show that the generating function for $(-1)^n \pi_{n,n-i}^{l,k}$ coincides with the generating function of $\pi_{n,i}^{k,l}$,

$$\begin{aligned} g(u, w, y, x) &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i (-1)^n \pi_{n,n-i}^{l,k} \\ &= \sum_{l=0}^{\infty} \frac{w^l}{l!} \sum_{k=0}^{\infty} \frac{u^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! (-y)^n \sum_{i=0}^n x^{n-i} \pi_{n,i}^{k,l} \\ &= h(w, u, -xy, 1/x) \\ &= \frac{\exp((-uxy + wy)/2)}{1 - (w + u)} \frac{\cosh(-wxy/2) \cosh(-uy/2)}{\cosh(-yx(u + w)(1 + 1/x)/2)} \\ &= \frac{\exp((wy - uxy)/2)}{1 - (w + u)} \frac{\cosh(wxy/2) \cosh(uy/2)}{\cosh(yx(u + w)(1 + 1/x)/2)} \\ &= h(u, w, y, x). \end{aligned}$$

□

4 \mathfrak{F} as a Lie algebra

Since \mathfrak{F} is an associative algebra, it is immediately obvious that it is also a Lie algebra with the commutator as a canonical Lie product. From (3.1), we know that the commutator can be expanded to

$$[\langle x \rangle_k, \langle y \rangle_l] = \sum_{n=0}^{k+l} \sum_{i=0}^n \mu_{n,i}^{k,l} \langle D^i x D^{n-i} y \rangle_{k+l-n}, \quad (4.1)$$

where $\mu_{n,i}^{k,l} = \pi_{n,i}^{k,l} - \pi_{n,n-i}^{l,k}$. However, the structure of \mathfrak{F} turns out to be more interesting than this.

Theorem 5 *\mathfrak{F} is a Lie algebra where commutators can be solved explicitly using the rule*

$$[\langle x \rangle_k, \langle y \rangle_l] = \sum_{n=0}^{\frac{k+l-1}{2}} \sum_{i=0}^{2n+1} \lambda_{n,i}^{k,l} \langle D^i x D^{2n+1-i} y \rangle_{k+l-2n-1}, \quad (4.2)$$

where $\lambda_{n,i}^{k,l} = 2\pi_{2n+1,i}^{k,l} \in \mathbb{Q}$.

Proof The even indexed coefficients $\mu_{2n,i}^{k,l}$ in (4.1) vanish due to Lemma 4, $\mu_{2n,i}^{k,l} = \pi_{2n,i}^{k,l} - (-1)^{2n} \pi_{2n,i}^{l,k} = 0$, while $\mu_{2n+1,i}^{k,l} = \pi_{2n+1,i}^{k,l} - (-1)^{2n+1} \pi_{2n+1,i}^{l,k} = 2\pi_{2n+1,i}^{k,l}$. We conveniently rename $\mu_{2n+1,i}^{k,l}$ as $\lambda_{n,i}^{k,l}$. \square

We now look at some interesting properties of this Lie algebra. Consider the linear spaces,

$$\mathfrak{G}_n = \bigoplus_{k \leq n} \mathfrak{F}_k.$$

We note that, as a consequence of Theorem 5, we have a natural filtration,

$$[\mathfrak{G}_k, \mathfrak{G}_l] \subseteq \mathfrak{G}_{k+l-1}.$$

This property of *height reduction* is evident in (4.2) where the largest index $k + l - 1$ occurs for $s = 0$. We define the height of a term in \mathfrak{G}_n as,

$$\text{ht} \left(\sum_{k=0}^n \langle x_k \rangle_k \right) = n,$$

if x_n is non-zero. We define the height of 0 as -1 , this being the only term with negative height. Thus $\text{ht}(\langle x \rangle_k) = k$ and, as a consequence of (4.2),

$$\text{ht}([\langle x \rangle_k, \langle y \rangle_l]) \leq k + l - 1. \quad (4.3)$$

Corollary 6 (Height Reduction) *For any commutator C featuring the letters $a_i \in \mathfrak{G}_{k_i}$, $i = 1, \dots, n$,*

$$\text{ht}(C(a_1, \dots, a_n)) \leq \sum_{i=1}^n k_i - n + 1. \quad (4.4)$$

Another useful property of this Lie algebra is that it doesn't *mix* terms of certain forms. We define

$$\mathfrak{e} = \bigoplus_{k \geq 0} \mathfrak{F}_{2k}, \quad \mathfrak{o} = \bigoplus_{k \geq 0} \mathfrak{F}_{2k+1},$$

so that $\mathfrak{F} = \mathfrak{o} \oplus \mathfrak{e}$. The following relations are evident from (4.2),

$$\begin{aligned} [\mathfrak{e}, \mathfrak{e}] &\subseteq \mathfrak{o}, \quad [\mathfrak{o}, \mathfrak{o}] \subseteq \mathfrak{o}, \\ [\mathfrak{e}, \mathfrak{o}] &\subseteq \mathfrak{e}, \quad [\mathfrak{o}, \mathfrak{e}] \subseteq \mathfrak{e}. \end{aligned} \tag{4.5}$$

Thus, if each letter a_i is either in \mathfrak{e} or in \mathfrak{o} , the commutator $C(a_i, \dots, a_n)$ is either in \mathfrak{e} or \mathfrak{o} and does not mix terms from the two. Moreover, commutators with the same letters will fall in the same space, \mathfrak{e} or \mathfrak{o} . This has implications for certain numerical methods where this property translates into a matrix being either symmetric or skew-symmetric but not a mix of the two which would have resulted in unfavourable structures.

The property (4.5), in fact, says that \mathfrak{F} is a \mathbb{Z}_2 -graded Lie algebra with the degree 0 component \mathfrak{o} and degree 1 component \mathfrak{e} . The linear space \mathfrak{o} is a Lie algebra in its own right, while structures of the form \mathfrak{e} are also called Lie triple systems which are closed under double commutation:

$$[\mathfrak{e}, [\mathfrak{e}, \mathfrak{e}]] \subseteq \mathfrak{e}.$$

These notions are closely related to Lie groups and symmetric spaces which have found applications in linear algebra methods such as the generalised polar decomposition (Munthe-Kaas, Quispel & Zanna 2001).

As anticipated, there is a close relation between \mathfrak{F} and the *free Lie algebra* generated by \mathcal{C} and polynomials in d .

Lemma 7 *The free Lie algebra generated by \mathcal{C} and $\mathcal{P}(d)$ (the ring of polynomials in d with constant coefficients) is contained in \mathfrak{F} ,*

$$\mathfrak{g} := \text{FLA}(\mathcal{C} \cup \mathcal{P}(d)) \subseteq \mathfrak{F}.$$

The two are identical if an inverse of the mapping $D = \text{ad}_d : \mathcal{C} \rightarrow \mathcal{C}$ exists.

Proof The containment is not difficult to prove. Since $d^k = \langle 1 \rangle_k$, it is contained in \mathfrak{F}_k and therefore $\mathcal{P}(d) \subseteq \mathfrak{F}$. The algebra \mathcal{C} is also contained in \mathfrak{F} since every $x \in \mathcal{C}$ can be written in the form $x = \langle x \rangle_0$. The free Lie algebra \mathfrak{g} generated by $\mathcal{C} \cup \mathcal{P}(d)$ is the intersection of all Lie algebras containing \mathcal{C} and $\mathcal{P}(d)$ and is therefore contained in the Lie algebra \mathfrak{F} .

The two algebras \mathfrak{g} and \mathfrak{F} are identical if inverse of the map D exists. Any term in \mathfrak{F}_0 is of the form $\langle y \rangle_0$ for some $y \in \mathcal{C}$ and therefore trivially resides in \mathfrak{g} . Take any $x \in \mathcal{C}$ and note that $d^2 \in \mathcal{P}(d)$. These are both contained in \mathfrak{g} and by definition of being a Lie algebra $[d^2, x] = [\langle 1 \rangle_2, \langle x \rangle_0] = \lambda_{0,0}^{2,0} \langle D^0(1)D^1(x) \rangle_1 + \lambda_{0,1}^{2,0} \langle D^1(1)D^0(x) \rangle_1 = 2 \langle D(x) \rangle_1$ resides in \mathfrak{g} as well. Consequently, any $\langle y \rangle_1 \in \mathfrak{F}_1$ can be expressed as $\frac{1}{2}[d^2, D^{-1}(y)]$ so long as the inverse of D exists. Thus we have $\mathfrak{F}_0, \mathfrak{F}_1 \subseteq \mathfrak{g}$. These two cases form the base case of our induction argument.

Assume that for all $k < n$, \mathfrak{F}_{2k} and \mathfrak{F}_{2k+1} are contained in \mathfrak{g} . For completing the inductive proof, consider

$$\begin{aligned} [d^{2n+1}, x] &= [\langle 1 \rangle_{2n+1}, \langle x \rangle_0] = \sum_{s=0}^n \sum_{i=0}^{2s+1} \lambda_{s,i}^{2n+1,0} \langle D^i(1) D^{2s+1-i}(x) \rangle_{2n-2s} \\ &= \sum_{s=0}^n \lambda_{s,0}^{2n+1,0} \langle D^{2s+1}(x) \rangle_{2n-2s} \\ &= \lambda_{0,0}^{2n+1,0} \langle D(x) \rangle_{2n} + \sum_{s=1}^n \lambda_{s,0}^{2n+1,0} \langle D^{2s+1}(x) \rangle_{2n-2s}. \end{aligned}$$

Setting $x = D^{-1}(y) \in \mathcal{C}$,

$$\langle y \rangle_{2n} = \frac{1}{\lambda_{0,0}^{2n+1,0}} [d^{2n+1}, D^{-1}(y)] - \sum_{s=1}^n \frac{\lambda_{s,0}^{2n+1,0}}{\lambda_{0,0}^{2n+1,0}} \langle D^{2s}(y) \rangle_{2n-2s}.$$

The first term on the right hand side is in \mathfrak{g} since $d^{2n+1} \in \mathcal{P}(d)$, while the terms in the summation fall in \mathfrak{g} due to the induction hypothesis. Thus, we find that $\langle y \rangle_{2n} \in \mathfrak{g}$ for any $y \in \mathcal{C}$. A similar proof shows that $\langle y \rangle_{2n+1}$ also resides in \mathfrak{g} , proving that $\mathfrak{F}_k \subseteq \mathfrak{g}$ for every k . Since \mathfrak{g} is also a linear space, the direct sum of these spaces, $\mathfrak{F} = \bigoplus_{k \in \mathbb{Z}_0^+} \mathfrak{F}_k$, is also contained in it. This completes our proof, $\mathfrak{g} = \mathfrak{F}$. \square

We note that, in general, the free Lie algebra \mathfrak{g} is not a subalgebra of the Lie idealiser \mathcal{I} since d^2 need not be in \mathcal{I} for every $d \in \mathcal{I}$.

Of less immediate and practical interest to us is the fact that \mathfrak{F} is also a Jordan algebra. This follows directly and trivially from the fact that it is an associative algebra (Lemma 1). More interestingly, it is also a \mathbb{Z}_2 -graded Jordan algebra since, along similar lines to (4.1),

$$\langle x \rangle_k \bullet \langle y \rangle_l = \sum_{n=0}^{k+l} \sum_{i=0}^n \gamma_{n,i}^{k,l} \langle D^i x D^{n-i} y \rangle_{k+l-n}, \quad (4.6)$$

where $\gamma_{n,i}^{k,l} = (\pi_{n,i}^{k,l} + \pi_{n,n-i}^{l,k})/2$ vanishes for odd values of n due to Lemma 4, while $\gamma_{2n,i}^{k,l} = \pi_{2n,i}^{k,l}$ survives. Corresponding observations about the *free Jordan algebra* generated by \mathcal{C} and $\mathcal{P}(d)$ can also be made along similar lines, however a property corresponding to height reduction of Corollary 6 does not follow.

5 Some applications

The applications of the algebras of Section 4 in solving partial differential equations such as the time-dependent Schrödinger equation will arise from treating function spaces as commutative algebras and linear differential operators as elements of the associative algebra of endomorphisms. With this motivation, we will restrict our attention to cases where the commutative algebra \mathcal{C} is isomorphic to a function space.

Let \mathcal{S} be a commutative algebra isomorphic to \mathcal{C} and Θ be the isomorphism between them. Anticipating the case where \mathcal{S} is a function space, we will use f, g, h to denote its elements. For $f \in \mathcal{S}$, we use the notation Θ_f and $\Theta(f)$ interchangeably for the corresponding element in \mathcal{C} .

For an element $d \in \mathcal{I}$ in the Lie idealiser of \mathcal{C} , $D = \text{ad}_d$ is a derivation on \mathcal{C} . For any $f \in \mathcal{S}$, $D(\Theta_f) = [d, \Theta_f] \in \mathcal{C}$ so that $\tilde{d}(f) := \Theta^{-1}(D(\Theta_f))$ is an element in \mathcal{S} . Since D is a derivation on \mathcal{C} and \mathcal{S} is isomorphic to \mathcal{C} , \tilde{d} is a derivation on \mathcal{S} . We say that \tilde{d} is a derivation induced by d . The element of the Lie idealiser that we will require in the following sections will be the differential operator $d = \partial_x$, which will end up coinciding with the induced derivation \tilde{d} .

Our first example of Θ will be \mathcal{M} , the left multiplication map, which we encounter while solving the TDSE. In the case of the Wigner equation, the isomorphism Θ is more complicated and maps functions to pseudo-differential operators.

5.1 Semiclassical time-dependent Schrödinger equation (TDSE)

The linear time-dependent Schrödinger equation (TDSE),

$$i\hbar\partial_t u(x, t) = -\frac{\hbar^2}{2m}\partial_x^2 u(x, t) + V(x)u(x, t), \quad x \in \mathbb{R}, \quad t \geq 0, \quad u(x, 0) = u_0(x),$$

is a fundamental equation of quantum mechanics. The reduced Planck's constant, $\hbar \approx 1.054 \times 10^{-34} \text{ J} \cdot \text{s}$, is a truly minute number which would typically result in considerable difficulties as far as numerical solutions are concerned. However, when working in atomic units, we can re-scale $\hbar = 1$. In these units the mass of an electron is 1. The typical length scales in these units are 10^{-11} m , while typical time and mass scales are 10^{-17} s and 10^{-30} kg , respectively. Consequently, working in the atomic units restricts us to extremely small spatio-temporal windows.

When computation for any larger spatio-temporal windows is required, we arrive at the TDSE under the semi-classical scaling,

$$i\varepsilon\partial_t u(x, t) = -\varepsilon^2\partial_x^2 u(x, t) + V(x)u(x, t), \quad x \in \mathbb{R}, \quad t \geq 0, \quad u(x, 0) = u_0(x). \quad (5.1)$$

The *semiclassical* parameter ε plays a role similar to the reduced Planck's constant. Although it is very small, $0 < \varepsilon \ll 1$, the semiclassical parameter is considerably larger than the Planck's constant and the range $10^{-8} \leq \varepsilon \leq 10^{-2}$ isn't unrealistic.

The equation (5.1) also arises out of the *Born–Oppenheimer approximation* of the molecular Schrödinger equation. In this case it describes the evolution of the nuclei in the electric potential, $V(x)$. Working in the atomic units, ε is the square root of the ratio of the mass of an electron and the heaviest nucleus. When larger spatio-temporal windows are required, ε can decrease further in size as before.

Typical computational methods for solving this equation commence with imposition of periodic boundaries at ± 1 followed by a discretisation of space. The infinite dimensional operators ∂_x and V are replaced by matrices \mathcal{K} and \mathcal{D}_V , respectively. The TDSE is now replaced by a system of ODEs,

$$\partial_t \mathbf{u}(t) = (i\varepsilon\mathcal{K}^2 - i\varepsilon^{-1}\mathcal{D}_V) \mathbf{u}(t), \quad t \geq 0, \quad \mathbf{u}(0) = \mathbf{u}_0,$$

where \mathbf{u} is the vector of values at grid points. The exact solution of this equation can be written as

$$\mathbf{u}(t) = \exp(it\varepsilon\mathcal{K}^2 - it\varepsilon^{-1}\mathcal{D}_V) \mathbf{u}_0.$$

Methods for evaluating the exact matrix exponential for this matrix are prohibitively expensive and one resorts to splitting methods such as the Trotter splitting,

$$\exp(it\varepsilon\mathcal{K}^2 - it\varepsilon^{-1}\mathcal{D}_V) = \exp(it\varepsilon\mathcal{K}^2) \exp(-it\varepsilon^{-1}\mathcal{D}_V) + \mathcal{O}(t^2),$$

or the Strang splitting,

$$\exp(it\varepsilon\mathcal{K}^2 - it\varepsilon^{-1}\mathcal{D}_V) = \exp\left(\frac{1}{2}it\varepsilon\mathcal{K}^2\right) \exp(-it\varepsilon^{-1}\mathcal{D}_V) \exp\left(\frac{1}{2}it\varepsilon\mathcal{K}^2\right) + \mathcal{O}(t^3). \quad (5.2)$$

Methods with higher-order accuracy than the Strang splitting can be created using Yoshida composition (Yoshida 1990). However the number of exponentials grows exponentially with the order desired. Other methods are based on the *symmetric Baker-Campbell-Hausdorff* (sBCH) formula (Dynkin 1947, Casas & Murua 2009),

$$\exp\left(\frac{1}{2}tA\right) \exp(tB) \exp\left(\frac{1}{2}tA\right) = \exp(\text{sBCH}(tA, tB)),$$

where

$$\text{sBCH}(tA, tB) = t(A + B) - t^3\left(\frac{1}{24}[[B, A], A] + \frac{1}{12}[[B, A], B]\right) + \mathcal{O}(t^5). \quad (5.3)$$

These require the evaluation of nested commutators of the matrices \mathcal{K}^2 and \mathcal{D}_V such as $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$, which is prohibitive and therefore typically avoided.

In (Bader et al. 2014) we develop arbitrarily high order methods by working directly within the free Lie algebra generated by the undiscretised operators ∂_x^2 and V . The cost of these methods grows quadratically in contrast to Yoshida methods. Additionally, these methods are commutator-free and preserve unitary evolution of the wave function—a fundamental principle of quantum mechanics. In this section we discover how the remarkable properties of symmetric Zassenhaus methods of (Bader et al. 2014) arise from the structural properties of (a special case of) the Lie algebra \mathfrak{F} .

5.1.1 Notation

Let $\mathcal{S} \subseteq \mathcal{T}$ be real and complex valued function spaces, respectively. Endomorphisms on \mathcal{T} form an associative algebra $\mathcal{A} = (\text{End}(\mathcal{T}), \circ, +)$ where \circ is operator composition. For convenience, we consider $\mathcal{S} = C_p^\infty([-1, 1]; \mathbb{R})$ and $\mathcal{T} = C_p^\infty([-1, 1]; \mathbb{C})$, the space of smooth periodic functions over $[-1, 1]$ with values in \mathbb{R} and \mathbb{C} , respectively. Let $M : \mathcal{S} \rightarrow \text{End}(\mathcal{T})$ be the left multiplication map,

$$M(f)(g) = fg \in \mathcal{T}, \quad f \in \mathcal{S}, g \in \mathcal{T}.$$

The image of \mathcal{S} under M ,

$$\mathcal{C} = M(\mathcal{S}),$$

is a commutative algebra of multiplication operators with M acting as an isomorphism between \mathcal{S} and \mathcal{C} . We also write M_f to denote the operator $M(f)$ whose action is that of multiplying by f . The partial differentiation operator ∂_x is an element of the Lie idealiser of \mathcal{C} in the associative operator algebra $(\text{End}(\mathcal{T}), \circ, +)$,

$$[\partial_x, M_f] = \partial_x \circ M_f - M_f \circ \partial_x = M_{\partial_x f} + M_f \circ \partial_x - M_f \circ \partial_x = M_{\partial_x f} \in \mathcal{C},$$

where the commutator is the canonical Lie product as usual. Following our terminology, the operator $\partial_x \in \text{End}(\mathcal{T})$ induces the derivation $\tilde{\partial}_x = \partial_x$ on \mathcal{S} . Here the induced derivation $\tilde{\partial}$ overlaps with the element d , but this need not be the case in general. To be very precise, $\tilde{\partial}_x$ is an operator on \mathcal{S} while ∂_x is an operator on \mathcal{T} , but the two coincide on \mathcal{S} .

Lemma 7 directly allows us to conclude that

$$\text{FLA}(\mathcal{M}(\mathcal{S}) \cup \mathcal{P}(\partial_x)) = \mathfrak{F}^{\partial_x} = \bigoplus_{k \in \mathbb{Z}_0^+} \{\langle \mathcal{M}_f \rangle_k^{\partial_x} : f \in \mathcal{S}\}.$$

Here D^{-1} is isomorphic to \tilde{d}^{-1} which is the inverse of differentiation on the space \mathcal{S} . Assuming \mathcal{S} is closed under integration, therefore, an appropriate inverse of D exists and the free Lie algebra of multiplicative operators $\mathcal{M}(\mathcal{S})$ and polynomials (with constant coefficients) in the differential operators, $\mathcal{P}(\partial_x)$, is characterised completely by $\mathfrak{F}^{\partial_x}$.

For convenience, we abuse notation and write

$$\langle f \rangle_k = \langle M_f \rangle_k^{\partial_x} = \frac{1}{2} \left(M_f \circ \partial_x^k + \partial_x^k \circ M_f \right).$$

In this notation

$$\langle 1_{\mathcal{S}} \rangle_k = \partial_x^k, \quad \langle f \rangle_0 = M_f,$$

where $1_{\mathcal{S}}$ is the constant function over \mathcal{S} with value 1.

5.1.2 Consequences for exponential splitting schemes

The semiclassical TDSE (5.1),

$$\partial_t u = i\varepsilon \partial_x^2 u - i\varepsilon^{-1} V u,$$

can be written in the form

$$\partial_t u = i(\varepsilon \langle 1 \rangle_2 - \varepsilon^{-1} \langle V \rangle_0) u,$$

where the unit in $\langle 1 \rangle_2$ is understood to be $1_{\mathcal{S}}$. We can find the solution of this equation by formally exponentiating the Hamiltonian without discretisation,

$$u(t) = \exp(it\varepsilon \langle 1 \rangle_2 - it\varepsilon^{-1} \langle V \rangle_0) u(0).$$

We may now perform splittings directly on this undiscretised Hamiltonian.

Methods which featured nested matrix commutators such as $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$ now feature corresponding commutators of operators $\langle 1 \rangle_2$ and $\langle V \rangle_0$, such as $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2]$. Since the two operators reside in our Lie algebra,

$$\langle 1 \rangle_2, \langle V \rangle_0 \in \mathfrak{F} = \bigoplus_{k \in \mathbb{Z}_0^+} \{ \langle f \rangle_k : f \in \mathcal{S} \},$$

so do all of their commutators. Moreover, as a consequence of Theorem 5, commutators in this Lie algebra can be solved explicitly,

$$[\langle f \rangle_k, \langle g \rangle_l] = \sum_{n=0}^{\frac{k+l-1}{2}} \sum_{i=0}^{2n+1} \lambda_{n,i}^{k,l} \langle (\partial_x^i f)(\partial_x^{2n+1-i} g) \rangle_{k+l-2n-1}, \quad (5.4)$$

where λ s are the same as before. This allows us to design commutator-free methods by explicitly working out nested commutators such as $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2]$ and $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle V \rangle_0]$ appearing in the sBCH of $it\varepsilon \langle 1 \rangle_2$ and $-it\varepsilon^{-1} \langle V \rangle_0$ using the rules,

$$\begin{aligned} [\langle f \rangle_2, \langle g \rangle_1] &= -\langle (\partial_x f)(\partial_x^2 g) + \frac{1}{2} f(\partial_x^3 g) \rangle_0 + \langle 2f(\partial_x g) - (\partial_x f)g \rangle_2, \\ [\langle f \rangle_2, \langle g \rangle_0] &= 2 \langle f(\partial_x g) \rangle_1, \\ [\langle f \rangle_1, \langle g \rangle_0] &= \langle f(\partial_x g) \rangle_0, \end{aligned}$$

which can be read off Table 2 in Appendix A. Consequently,

$$\begin{aligned} [\langle V \rangle_0, \langle 1 \rangle_2] &= -2 \langle \partial_x V \rangle_1, \\ [[\langle V \rangle_0, \langle 1 \rangle_2], \langle V \rangle_0] &= -2 \langle (\partial_x V)^2 \rangle_0, \\ [[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2] &= -\langle \partial_x^4 V \rangle_0 + 4 \langle \partial_x^2 V \rangle_2. \end{aligned}$$

The sBCH of $it\varepsilon \langle 1 \rangle_2$ and $-it\varepsilon^{-1} \langle V \rangle_0$ up to $\mathcal{O}(t^5)$, for instance, is

$$it\varepsilon \langle 1 \rangle_2 - it\varepsilon^{-1} \langle V \rangle_0 - \frac{1}{6}it^3\varepsilon \langle \partial_x^2 V \rangle_2 + \frac{1}{24}it^3\varepsilon \langle \partial_x^4 V \rangle_0 - \frac{1}{6}t^3\varepsilon^{-1} \langle (\partial_x V)^2 \rangle_0.$$

Another favourable consequence of Theorem 5 is that all terms in the sBCH of $it\varepsilon \langle 1 \rangle_2$ and $-it\varepsilon^{-1} \langle V \rangle_0$ reside in

$$\mathfrak{H} = \bigoplus_{k \in \mathbb{Z}_0^+} \{i^{k+1} \langle f \rangle_k : f \in \mathcal{S}\},$$

whose elements are skew-Hermitian operators. Upon discretisation using pseudospectral methods, these are replaced by

$$i^{k+1} \langle f \rangle_k \rightsquigarrow \frac{1}{2}i^{k+1}(\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f),$$

where \mathcal{D}_f is a diagonal matrix and \mathcal{K} a skew-symmetric circulant differentiation matrix. Consequently, the discretised forms of elements of \mathfrak{H} are skew-Hermitian matrices, the exponentials of which are unitary matrices. Zassenhaus splittings (Bader et al. 2014) feature exponentials of terms in \mathfrak{H} , thereby guaranteeing unitary evolution of the wavefunction and resulting in unconditional stability of these numerical methods.

It is well known that solutions of the semiclassical TDSE (5.1)—irrespective of smooth initial conditions—develop oscillations of frequency $\mathcal{O}(\varepsilon^{-1})$ in both space and time (Jin, Markowich & Sparber 2011). These oscillations make computational solutions very costly since one typically needs $M = \mathcal{O}(\varepsilon^{-1})$ grid points to resolve spatial oscillations and time steps of size $\mathcal{O}(\varepsilon)$ to resolve temporal oscillations.

The spectral radius of the differentiation matrix \mathcal{K} scales as $\mathcal{O}(M)$, i.e. $\mathcal{O}(\varepsilon^{-1})$. Keeping eventual discretisation in mind and noting that

$$\langle f \rangle_k \rightsquigarrow \frac{1}{2}(\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f) = \mathcal{O}(\varepsilon^{-k}),$$

we use the shorthand $\langle f \rangle_k = \mathcal{O}(\varepsilon^{-k})$. It is now easy to see how the height of a term $a \in \mathfrak{G}_n$ acts as a proxy for the spectral radius of the eventual discretisation,

$$a \in \mathfrak{G}_n \iff \text{ht}(a) \leq n \implies a = \mathcal{O}(\varepsilon^{-n}).$$

Since the matrix \mathcal{K}^2 scales as $\mathcal{O}(\varepsilon^{-2})$, one would expect the commutator $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$ to scale as $\mathcal{O}(\varepsilon^{-4})$. The undiscretised commutator $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2]$, on the other hand, scales as $\mathcal{O}(\varepsilon^{-2})$ thanks to the property of height reduction (Corollary 6).

Corollary 8 Commutators of $A = i\varepsilon \langle 1 \rangle_2$ and $B = i\varepsilon^{-1} \langle V \rangle_0$ are of size $\mathcal{O}(\varepsilon^{-1})$.

Proof Consider a grade n commutator C of $\langle 1 \rangle_2$ and $\langle V \rangle_0$, featuring k occurrences of the letter $\langle 1 \rangle_2$ and $n - k$ occurrences of $\langle V \rangle_0$. Since $\langle 1 \rangle_2 \in \mathfrak{G}_2$ and $\langle V \rangle_0 \in \mathfrak{G}_0$, using Corollary 6,

$$\text{ht}(C) \leq 2k - n + 1.$$

Thus C is $\mathcal{O}(\varepsilon^{-2k+n-1})$. However, the corresponding commutator of A and B is scaled by k occurrences of ε and $n - k$ occurrences of ε^{-1} , bringing its size to $\mathcal{O}(\varepsilon^{-1})$. \square

When we scale the time step as $t = \mathcal{O}(\varepsilon^\sigma)$ for some $\sigma > 0$, a grade n commutator of $A = it\varepsilon \langle 1 \rangle_2$ and $B = it\varepsilon^{-1} \langle V \rangle_0$, scales as $\mathcal{O}(\varepsilon^{n\sigma-1})$ in contrast to the $\mathcal{O}(\varepsilon^{n(\sigma-1)})$ scaling seen when working with commutators of matrices. Whereas the latter requires $\sigma > 1$ for

convergence of sBCH (which corresponds to very small time steps), our approach does not have such restrictions and allows convergence for extremely large time steps. With time steps of size $\mathcal{O}(\sqrt{\varepsilon})$, which corresponds to $\sigma = 1/2$, for instance, grade n commutators are $\mathcal{O}(\varepsilon^{n/2-1})$ and the sBCH series easily converges.

The property of height reduction underlies the asymptotic splitting of the symmetric Zassenhaus kind where it results in inexpensive exponentiations via Lanczos iterations (Bader et al. 2014). The advantages of height reduction also extend to error analysis for existing methods such as Yoshida splittings since commutators discarded in these methods can be assumed to be undiscretised and analysed appropriately. An order six Yoshida splitting discards all commutators of grade seven and higher, for instance, thereby committing an error of $\mathcal{O}(\varepsilon^{7\sigma-1})$.

Despite this, the quadratic cost of the Zassenhaus splitting, which results from a systematic exploitation of the structures of \mathfrak{F} and the property of height reduction, trumps the exponential cost of Yoshida.

Theorem 9 *Cost of Zassenhaus splittings grows quadratically in the order desired.*

Proof Zassenhaus splittings (Bader et al. 2014) recursively utilise the sBCH formula featuring odd grade commutators of $it\varepsilon\langle 1 \rangle_2$ and $-it\varepsilon^{-1}\langle V \rangle_0$. Due to Corollary 8, any grade n commutator of $it\varepsilon\langle 1 \rangle_2$ and $-it\varepsilon^{-1}\langle V \rangle_0$ is $\mathcal{O}(\varepsilon^{n\sigma-1})$. Splittings with $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ error, therefore, require commutators of grades up to $2n+1$. Such a splitting has the form,

$$e^{it\varepsilon\langle 1 \rangle_2 - \varepsilon^{-1}\langle V \rangle_0} = e^{\frac{1}{2}W^{[0]}} \dots e^{\frac{1}{2}W^{[n]}} e^{W^{[n+1]}} e^{\frac{1}{2}W^{[n]}} \dots e^{\frac{1}{2}W^{[0]}} + \mathcal{O}(\varepsilon^{(2n+3)\sigma-1}),$$

where $W^{[0]} = it\varepsilon\langle 1 \rangle_2$ and $W^{[1]} = -it\varepsilon^{-1}\langle V \rangle_0$, both of which are $\mathcal{O}(\varepsilon^{\sigma-1})$. For $k \geq 2$, $W^{[k]}$ arises from grade $2k-1$ commutators of $it\varepsilon\langle 1 \rangle_2$ and $-it\varepsilon^{-1}\langle V \rangle_0$, and is, therefore, of size $\mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$. We refer the reader to (Bader et al. 2014) for details of these splittings.

The first exponent, $W^{[0]}$, is discretised as a circulant matrix, $it\varepsilon\mathcal{K}^2$, whose exponential can be evaluated using two Fast Fourier Transforms (FFTs), each of which costs $\mathcal{O}(M \log M) = \mathcal{O}(\varepsilon^{-1} \log(\varepsilon^{-1}))$. The second exponent, $W^{[1]}$, is discretised as a diagonal matrix, $-it\varepsilon^{-1}\mathcal{D}_V$, and can be exponentiated directly.

Remaining exponents require Lanczos iterations. Since $W^{[k]}$ is $\mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$, we need $\left\lceil \frac{(2n+3)\sigma-1}{(2k-1)\sigma-1} \right\rceil$ iterations for approximating its exponential to $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ accuracy (Bader et al. 2014). Each of these iterations require the evaluation of matrix-vector products of the form $W^{[k]}u$.

Since $W^{[k]}$ arises from grade $2k-1$ commutators of the terms $it\varepsilon\langle 1 \rangle_2 \in \mathfrak{G}_2 \cap \mathfrak{e}$ and $-it\varepsilon^{-1}\langle V \rangle_0 \in \mathfrak{G}_0 \cap \mathfrak{e}$, it resides in the intersection of \mathfrak{G}_{2k-2} (at most) and \mathfrak{e} . Thus, it consists of terms of the form $\langle f_0 \rangle_0, \langle f_2 \rangle_2, \langle f_4 \rangle_4, \dots, \langle f_{2k-2} \rangle_{2k-2}$. To evaluate $W^{[k]}u$, we need to evaluate $\langle f_0 \rangle_0 u, \langle f_2 \rangle_2 u, \langle f_4 \rangle_4 u, \dots, \langle f_{2k-2} \rangle_{2k-2} u$. The first of these, $\langle f_0 \rangle_0 u$, is a pointwise product, while the rest require four FFTs each since $\langle f \rangle_k$ is discretised as $(\mathcal{D}_f \mathcal{K}^2 + \mathcal{K}^2 \mathcal{D}_f)/2$.

The cost of the splitting is dominated by the cost of FFT operations, each of which requires $\mathcal{O}(\varepsilon^{-1} \log(\varepsilon^{-1}))$ operations. Since each evaluation of $W^{[k]}u$ requires $4(k-1)$ FFTs, the number of FFTs required per time step of the splitting scheme is

$$C_\sigma(n) = 4 + 2 \sum_{k=2}^n 4(k-1) \left\lceil \frac{(2n+3)\sigma-1}{(2k-1)\sigma-1} \right\rceil + 4n \left\lceil \frac{(2n+3)\sigma-1}{(2n+1)\sigma-1} \right\rceil,$$

which grows quadratically. Under $\sigma = 1$, for instance, this grows as $\sim 12n^2$,

$$\begin{aligned} C_1(n) &= 4 + 2 \sum_{k=2}^n 4(k-1) \left\lceil \frac{2n+2}{2k-2} \right\rceil + 4n \left\lceil \frac{2n+2}{2n} \right\rceil \\ &\leq 4 + 8 \sum_{k=2}^n (k-1) \left(\frac{2n+2}{2k-2} + 1 \right) + 8n \\ &= 4 + 8 \sum_{k=2}^n (n+k) + 8n = 12n^2 + 4n - 4. \end{aligned}$$

The overall cost for an $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ splitting is $\mathcal{O}(C_\sigma(n)\varepsilon^{-1} \log \varepsilon^{-1})$ per time step. We remark that careful choices can allow us to reduce the exact number of FFTs required. However, we are largely concerned with asymptotic growths here.

Since this proof does not assume a specific form of the exponents, $W^{[k]}_S$, and works solely on the basis of the structure of \mathfrak{G}_k and \mathfrak{e} , the odd graded nature of the sBCH formula and the height reduction of Corollary 6, the cost estimates derived here also apply directly to Magnus–Zassenhaus approaches of (Iserles et al. 2015, Bader et al. 2015) discussed in Section 5.2. \square

5.1.3 Some related algebras

The gain of powers of ε has also been observed by Gaim & Lasser (2014) using Moyal brackets in the phase space. However, the analysis is not as easily generalised as the algebraic approach here.

These ideas are closely related to the Weyl algebra (Dixmier 1968, Coutinho 1997) which is the universal enveloping algebra of the Lie algebra of Heisenberg groups. The univariate Weyl algebra can be written in the form

$$\mathfrak{W} = \left\{ \sum_{k=0}^n f_k(x) \partial_x^k : f_k \in \mathcal{P}(x), k \in \mathbb{Z}_0^+ \right\}.$$

These are special cases of a more general form

$$\mathfrak{S} = \left\{ \sum_{k=0}^n f_k(x) \partial_x^k : f_k \in \mathcal{S}, k \in \mathbb{Z}_0^+ \right\},$$

under the choice of polynomials in x , $\mathcal{P}(x)$, as the function space \mathcal{S} .

We note that \mathfrak{S} is contained within \mathfrak{F} as an associative algebra since $\mathcal{M}(f_k), \partial_x^k \in \mathfrak{F}$, while, in the other direction, $\mathfrak{F} \subseteq \mathfrak{S}$ is evident using the Leibniz rule. Thus, the two are identical as associative algebras. It is not difficult to prove that a height reduction phenomenon also holds for \mathfrak{S} .

However, there is a significant advantage to working in the symmetrised form of \mathfrak{F} since a \mathfrak{S} lacks a \mathbb{Z}_2 -graded structure. The \mathbb{Z}_2 grading on \mathfrak{F} and the symmetric nature of its elements ensures that elements of \mathfrak{H} are skew-Hermitian after discretisation. This proves crucial for devising stable numerical schemes such as the symmetric Zassenhaus schemes of (Bader et al. 2014) once we start utilising nested commutators. In contrast, lacking a \mathbb{Z}_2 grading, a clean separation of terms does not occur in the form \mathfrak{S} .

One could argue that any nested commutator of skew-Hermitian operators in \mathfrak{S} should be skew-Hermitian. This is indeed true prior to discretisation. However, unless some highly specialised differentiation matrices can be constructed, the discretised version of the commutator simplified in \mathfrak{S} possesses no such structure. What is more, even prior to discretisation we prefer to discard terms smaller than a certain size (while analysing in powers of ε). Working in \mathfrak{S} instead of \mathfrak{F} , it becomes difficult to discern which components of a term such as $\sum_{k=0}^n f_k(x) \partial_x^k$ can be discarded and which need to be kept despite their small size in order to preserve the skew-Hermiticity of the undiscretised operator.

5.2 Time-varying potentials in the TDSE

Consider the TDSE with time-varying electric potential, $V(x, t)$, written in the form

$$\partial_t u = \mathcal{A}(t)u,$$

where $\mathcal{A}(t) = i\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \langle V(t) \rangle_0 \in \mathfrak{H}$. The solution of such Lie group equations can be expressed using the Magnus expansion (Magnus 1954),

$$u(t) = e^{\Omega(t)} u(0),$$

where the infinite series $\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)$, is an element of the underlying Lie algebra \mathfrak{H} . The exponent $\Omega(t)$ satisfies the *dexpinv equation* (Iserles & Nørsett 1999),

$$\Omega'(t) = \text{dexp}_{\Omega(t)}^{-1}(\mathcal{A}(t)) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \text{ad}_{\Omega(t)}^k(\mathcal{A}(t)), \quad \Omega(0) = 0, \quad (5.5)$$

where B_k are the Bernoulli numbers. The solution of (5.5), originally proposed as an infinite series using Picard iterations (Magnus 1954), has been widely analysed in (Iserles & Nørsett 1999, Iserles, Munthe-Kaas, Nørsett & Zanna 2000, Blanes, Casas, Oteo & Ros 2009).

The graded version of the Magnus expansion (Iserles & Nørsett 1999) reads

$$\begin{aligned} \Omega(t) = & \int_0^t \mathcal{A}(\xi) d\xi - \frac{1}{2} \int_0^t \int_0^{\xi_1} [\mathcal{A}(\xi_2), \mathcal{A}(\xi_1)] d\xi \\ & + \frac{1}{12} \int_0^t \int_0^{\xi_1} \int_0^{\xi_1} [\mathcal{A}(\xi_2), [\mathcal{A}(\xi_3), \mathcal{A}(\xi_1)]] d\xi \\ & + \frac{1}{4} \int_0^t \int_0^{\xi_1} \int_0^{\xi_2} [[\mathcal{A}(\xi_3), \mathcal{A}(\xi_2)], \mathcal{A}(\xi_1)] d\xi + \dots \end{aligned}$$

Once again, since $\mathcal{A}(t) \in \mathfrak{H}$, all commutators can be simplified to commutator-free terms in \mathfrak{H} . This is in contrast to (Hochbruck & Lubich 2003) where the Magnus expansion features nested commutators of matrices.

Bounds for convergence of the Magnus expansion analysed in (Iserles & Nørsett 1999) were improved upon in (Moan & Niesen 2008). Applied directly to the semiclassical TDSE, these would impose a highly stringent restriction of $t = \mathcal{O}(\varepsilon^2)$ on the time step. Hochbruck & Lubich (2003) analysed the convergence of the Magnus expansion in the context of the TDSE, making significant improvements on the earlier and more general bounds. Nevertheless, under semiclassical scaling their analysis suggests a need for very small time steps of size $\mathcal{O}(\varepsilon)$ for convergence.

However, due to Corollary 8, a grade n term in the Magnus expansion of $\mathcal{A}(t)$ is $\mathcal{O}(\varepsilon^{n\sigma-1})$. Thus, asymptotically speaking in terms of ε , the terms in the expansion are decreasing in size

with increasing n for any $\sigma > 0$, so that convergence of the Magnus expansion also occurs for much larger time steps such as $t = \mathcal{O}(\sqrt{\varepsilon})$. This is a considerable improvement over earlier analysis.

Different versions of the Magnus expansion based on these observations have been successfully combined with Zassenhaus splittings to yield efficient methods for solving the TDSE with time-varying potentials (Iserles et al. 2015, Bader et al. 2015). The skew-Hermitian nature of elements of \mathfrak{H} once again leads to unitary evolution of our methods. As noted in Theorem 9, the cost of these methods also grows quadratically in the order desired. This should be contrasted against Yoshida based methods—instead of requiring 3 exponentials, the Strang splitting features 5 or more terms for the first non-trivial Magnus expansion and consequently the cost of Yoshida methods (which are derived by composing Strang splittings) grows as $\mathcal{O}(5^n)$ for an order- $2n$ accuracy. This is considerably more expensive than the $\mathcal{O}(3^n)$ cost for time-independent potentials.

5.3 A finite dimensional example

While the examples of the previous sections are based in infinite dimensional algebras and function spaces, it is also possible to construct finite dimensional examples of such structures. Consider the commutative subalgebra of $2n \times 2n$ matrices, $\mathcal{A} = M_{2n}(\mathbb{R})$,

$$\mathcal{C} = \left\{ \begin{pmatrix} aI_n & A \\ O_n & aI_n \end{pmatrix} : a \in \mathbb{R}, A \in M_n(\mathbb{R}) \right\},$$

where I_n is the $n \times n$ identity matrix and O_n is the $n \times n$ zero matrix. Consider any

$$d = \begin{pmatrix} E^{11} & E^{12} \\ O_n & E^{22} \end{pmatrix},$$

where $E^{ij} \in M_n(\mathbb{R})$. Then for every $x \in \mathcal{C}$,

$$[d, x] = \begin{pmatrix} O_n & E^{11}A - AE^{22} \\ O_n & O_n \end{pmatrix} \in \mathcal{C}.$$

Thus, d is in the Lie idealiser of \mathcal{C} , $D = ad_d$ is a derivation on \mathcal{C} , and, consequently,

$$\mathfrak{F}^d = \bigoplus_{k \in \mathbb{Z}_0^+} \{ \langle x \rangle_k^d = \frac{1}{2}(xd^k + d^kx) : x \in \mathcal{C} \}$$

is a \mathbb{Z}_2 -graded matrix Lie algebra.

Similar structures should be found in other commutative subalgebras of $M_n(\mathbb{R})$. Applications of these structural observations to linear algebra algorithms are yet to be explored.

5.4 Future work

1. Significantly relaxed bounds, improving on the bounds implied by (Suzuki 1977, Hochbruck & Lubich 2003, Moan & Niesen 2008, Casas, Murua & Nadinic 2012) for the convergence of the sBCH, Magnus and Zassenhaus expansions in the case of semiclassical TDSEs have been obtained due to the property of height reduction. Tighter bounds will require a careful analysis of the growth of the coefficients $\pi_{n,i}^{k,l}$ and derivatives of functions in (5.4).

2. The TDSEs analysed here are one-dimensional. Extending these ideas to $\partial_t u = i\varepsilon \Delta u - i\varepsilon^{-1}V(\mathbf{x})u$ will require an analysis of the free Lie algebra generated by polynomials in the commuting elements of the Lie idealiser, ∂_x, ∂_y and ∂_z .
3. The Wigner equation $\partial_t w(x, \xi, t) = -\xi \partial_x w(x, \xi, t) + \Theta^\varepsilon[\delta V] w(x, \xi, t)$, where $\Theta^\varepsilon[\delta V]$ is a pseudo-differential operator, features operators which are considerably more complicated (Jin et al. 2011). Yet, the Lie algebraic structure is similar to \mathfrak{F} with the role of the isomorphism played by Θ . In addition to ∂_x , multiplication by the frequency variable ξ is also an element of the Lie idealiser. While ∂_x induces the derivation ∂_x as usual, ξ induces the derivation $i\partial_y$.
4. Matrix-valued potentials appear in TDSEs when we need to consider multiple energy levels at once. Magnus expansion based methods such as (Hochbruck & Lubich 2003) are used once these potentials start featuring time-varying components (Kormann, Holmgren & Karlsson 2008).

Unfortunately, matrix-valued potentials do not directly fall into the framework of \mathfrak{F} as a Lie algebra. However suitable extensions for these contexts are being actively sought. Our initial findings suggest that a modified version of the height reduction rule holds and it might be possible to extend the significantly milder time step restrictions and favourable computational complexity of (Bader et al. 2014, Iserles et al. 2015, Bader et al. 2015) to these cases.

Acknowledgements

The author is grateful to Marjorie Batchelor, Kurusch Ebrahimi-Fard, Dimitar Grantcharov, Arieh Iserles, Caroline Lasser, Hans Munthe-Kaas and Antonella Zanna for guidance on algebraic notions, help with combinatorial techniques and for many helpful discussions.

References

- Bader, P., Iserles, A., Kropielnicka, K. & Singh, P. (2014), ‘Effective approximation for the semiclassical Schrödinger equation’, *Found. Comput. Math.* **14**, 689–720.
- Bader, P., Iserles, A., Kropielnicka, K. & Singh, P. (2015), Efficient methods for time-dependence in linear Schrödinger equation. Preprint available.
- Blanes, S., Casas, F., Oteo, J. A. & Ros, J. (2009), ‘The Magnus expansion and some of its applications’, *Phys. Rep.* **470**, 151–238.
- Casas, F. & Murua, A. (2009), ‘An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications’, *J. Math. Phys.* **50**, 033513–1–033513–23.
- Casas, F., Murua, A. & Nadinic, M. (2012), ‘Efficient computation of the Zassenhaus formula’, *Comp. Phys. Comm.* **183**, 2386–2391.
- Coutinho, S. C. (1997), ‘The many avatars of a simple algebra’, *Amer. Math. Monthly* **104**(7), 593–604.
- Dixmier, J. (1968), ‘Sur les algèbres de Weyl’, *Bulletin de la Société Mathématique de France* **96**, 209–242.

- Dynkin, E. B. (1947), ‘Evaluation of the coefficients of the Campbell-Hausdorff formula’, *Dokl. Akad. Nauk SSSR* **57**, 323–326.
- Gaim, W. & Lasser, C. (2014), ‘Corrections to Wigner type phase space methods’, *Nonlinearity* **27**(12), 2951.
- Hochbruck, M. & Lubich, C. (2003), ‘On magnus integrators for time-dependent Schrödinger equations’, *SIAM J. Numer. Anal.* **41**(3), 945–963.
- Iserles, A., Kropielnicka, K. & Singh, P. (2015), ‘On the discretisation of the semiclassical Schrödinger equation with time-dependent potential’. Technical Report NA2015/02.
- Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P. & Zanna, A. (2000), ‘Lie-group methods’, *Acta Numerica* **9**, 215–365.
- Iserles, A. & Nørsett, S. P. (1999), ‘On the solution of linear differential equations in Lie groups’, *Phil. Trans. R. Soc. A* **357**, 983–1019.
- Jin, S., Markowich, P. & Sparber, C. (2011), ‘Mathematical and computational methods for semiclassical Schrödinger equations’, *Acta Numerica* **20**, 121–209.
- Kormann, K., Holmgren, S. & Karlsson, H. O. (2008), ‘Accurate time propagation for the Schrödinger equation with an explicitly time-dependent Hamiltonian’, *J. Chem. Phys.* **128**(18).
- Magnus, W. (1954), ‘On the exponential solution of differential equations for a linear operator’, *Commun. Pure Appl. Math.* **7**, 649–673.
- Moan, P. C. & Niesen, J. (2008), ‘Convergence of the Magnus series’, *Found. Comput. Math.* **8**(3), 291–301.
- Munthe-Kaas, H. Z., Quispel, G. R. W. & Zanna, A. (2001), ‘Generalized polar decompositions on lie groups with involutive automorphisms’, *Found. Comput. Math.* **1**(3), 297–324.
- Suzuki, M. (1977), ‘On the convergence of exponential operators—the Zassenhaus formula, BCH formula and systematic approximations’, *Comm. Math. Phys.* **57**, 193–200.
- Yoshida, H. (1990), ‘Construction of higher order symplectic integrators’, *Phys. Lett.* **150**, 262–268.

A Tables of coefficients

The coefficients $\pi_{n,i}^{k,l}$ appearing in Lemma 1 for $k+l$ ranging between 1 and 4 are presented in Table 1. The case $\langle x \rangle_0 \cdot \langle y \rangle_0 = \langle xy \rangle_0$ is trivial and not listed. From Lemma 4 we know that $\pi_{n,i}^{k,l} = (-1)^n \pi_{n,n-i}^{l,k}$, so that specifying the rows (1, 3) and (0, 4), for instance, would be redundant.

In Table 2 we present the coefficients $\lambda_{n,i}^{k,l}$ appearing in Theorem 5 for $k+l$ ranging between 1 and 6, while noting that the relation $\lambda_{n,i}^{k,l} = -\lambda_{n,2n+1-i}^{l,k}$ makes redundant the need to specify coefficients when k and l exchange values. The values of the coefficients for $(k, l) = (0, 2)$ can be inferred from the row $(k, l) = (2, 0)$, for instance. Since $\lambda_{n,i}^{k,l} = 2\pi_{2n+1,i}^{k,l}$, the first eight rows can be read directly by doubling the corresponding rows in Table 1. Note that $[\langle x \rangle_0, \langle y \rangle_0] = 0$,

(k, l)	n	$\pi_{n,0}^{k,l}$	$\pi_{n,1}^{k,l}$	$\pi_{n,2}^{k,l}$	$\pi_{n,3}^{k,l}$	$\pi_{n,4}^{k,l}$
(1, 0)	0	1				
	1	1/2	0			
(2, 0)	0	1				
	1	1	0			
	2	0	-1/2	0		
(1, 1)	0	1				
	1	1/2	-1/2			
	2	-1/4	-3/4	-1/4		
(3, 0)	0	1				
	1	3/2	0			
	2	0	-3/2	0		
	3	-1/4	-3/4	0	0	
(2, 1)	0	1				
	1	1	-1/2			
	2	-1/2	2	-1/2		
	3	-1/4	-1/2	0	0	
(4, 0)	0	1				
	1	2	0			
	2	0	-3	0		
	3	-1	-3	0	0	
	4	0	1/2	3/2	1/2	0
(3, 1)	0	1				
	1	3/2	-1/2			
	2	-3/4	-15/4	-3/4		
	3	-1	-9/4	0	0	
	4	1/8	1	15/8	7/8	1/8
(2, 2)	0	1				
	1	1	-1			
	2	-1	-4	-1		
	3	-1/2	-1	1	1/2	
	4	1/4	5/4	9/4	5/4	1/4

Table 1: A table of the coefficients $\pi_{n,i}^{k,l}$, $n \in \{0, \dots, k+l\}$, $i \in \{0, \dots, n\}$, which appear in Lemma 1.

and the case $k+l=0$ doesn't merit a mention in the table. Using the row $(k, l) = (3, 2)$ we can compute

$$\begin{aligned}
 [\langle x \rangle_3, \langle y \rangle_2] &= \langle 3x Dy - 2Dxy \rangle_0 \\
 &\quad + \langle -(7/2)x D^3y - (15/2)DxD^2y + 3D^2xDy + (3/2)D^3xy \rangle_1 \\
 &\quad + \langle (3/4)x D^5y + 3DxD^4y + (7/2)D^2xD^3y - D^4xDy - (1/4)D^5xy \rangle_2.
 \end{aligned}$$

Similarly, in the context of the function space \mathcal{S} and differential operator $d = \partial_x$,

$$[\langle f \rangle_2, \langle g \rangle_1] = \langle -\frac{1}{2}f(\partial_x^3g) - (\partial_x f)(\partial_x^2g) \rangle_0 + \langle 2f(\partial_x g) - (\partial_x f)g \rangle_2$$

(k, l)	n	$\lambda_{n,0}^{k,l}$	$\lambda_{n,1}^{k,l}$	$\lambda_{n,2}^{k,l}$	$\lambda_{n,3}^{k,l}$	$\lambda_{n,4}^{k,l}$	$\lambda_{n,5}^{k,l}$
(1, 0)	0	1	0				
(2, 0)	0	2	0				
(1, 1)	0	1	-1				
(3, 0)	0	3	0				
	1	-1/2	-3/2	0	0		
(2, 1)	0	2	-1				
	1	-1/2	-1	0	0		
(4, 0)	0	4	0				
	1	-2	6	0	0		
(3, 1)	0	3	-1				
	1	-2	-9/2	0	0		
(2, 2)	0	2	-2				
	1	-1	-2	2	1		
(5, 0)	0	5	0				
	1	-5	-15	0	0		
	2	1	5	15/2	5/2	0	0
(4, 1)	0	4	-1				
	1	-5	-12	0	0		
	2	1	9/2	6	2	0	0
(3, 2)	0	3	-2				
	1	-7/2	-15/2	3	3/2		
	2	3/4	3	7/2	0	-1	-1/4
(6, 0)	0	6	0				
	1	-10	-30	0	0		
	2	6	30	45	15	0	0
(5, 1)	0	5	-1				
	1	-10	-25	0	0		
	2	6	55/2	75/2	25/2	0	0
(4, 2)	0	4	-2				
	1	-8	-18	4	2		
	2	5	21	26	4	-4	-1
(3, 3)	0	3	-3				
	1	-5	-21/2	21/2	5		
	2	3	12	21/2	-21/2	-12	-3

Table 2: A table of the coefficients $\lambda_{n,i}^{k,l}$, $n \in \{0, \dots, (k+l-1)/2\}$, $i \in \{0, \dots, 2n+1\}$, which appear in Theorem 5.

is found by substituting in (5.4) the coefficients from row $(k, l) = (2, 1)$. Note that these brackets are linear, $\langle \alpha x + \beta y \rangle_k = \alpha \langle x \rangle_k + \beta \langle y \rangle_k$.

B Proof of Lemma 2—explicit form of $\pi_{n,i}^{k,l}$

Let

$$S_{(a,p),(n,i)}^q = \begin{cases} \delta_{a,n}\delta_{p,i} + \binom{q-n}{a-n}\binom{a-n}{p-i} & a \geq n, p \geq i, \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.1})$$

The explicit form of π_s in (3.9) allows us to express $R_{a,p}^{k,l}$ in (3.6) as

$$\begin{aligned} R_{a,p}^{k,l} &= 2 \sum_{n=0}^{k+l} \sum_{i=0}^{k+l} S_{(a,p),(n,i)}^{k+l} \pi_{n,i}^{k,l} \\ &= \sum_{n=0}^{k+l} \sum_{i=0}^{k+l} S_{(a,p),(n,i)}^{k+l} \sum_{s=0}^{k+l} \sum_{j=0}^{k+l} A_{(n,i),(s,j)}^{k+l} L_{s,j}^{k,l} \\ &= \sum_{s=0}^{k+l} \sum_{j=0}^{k+l} \left[\sum_{n=0}^{k+l} \sum_{i=0}^{k+l} S_{(a,p),(n,i)}^{k+l} A_{(n,i),(s,j)}^{k+l} \right] L_{s,j}^{k,l}. \end{aligned}$$

To prove Lemma 2, we need to prove that A in (3.9) is such that $R_{a,p}^{k,l} = L_{a,p}^{k,l}$ is satisfied. This certainly holds if

$$\sum_{n=0}^q \sum_{i=0}^q S_{(a,p),(n,i)}^q A_{(n,i),(s,j)}^q = \delta_{a,s} \delta_{p,j}, \quad (\text{B.2})$$

holds for any q and $a, s, p, j \in \{0, \dots, q\}$. To prove this we note that S and A , and therefore their product SA , are lower triangular. Thus we may concern ourselves solely with the case $0 \leq s \leq a \leq q$ and $0 \leq j \leq p \leq q$. Denoting $(SA)_{(a,p),(s,j)}^q$ as T for brevity,

$$T = \sum_{n=0}^q \sum_{i=0}^q \left[\delta_{a,n} \delta_{p,i} + \binom{q-n}{a-n} \binom{a-n}{p-i} \right] \left[\delta_{n,s} \delta_{i,j} - \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j} \right]$$

can be separated into four parts,

$$\begin{aligned} T_1 &= \sum_{n=0}^q \sum_{i=0}^q \delta_{a,n} \delta_{p,i} \delta_{n,s} \delta_{i,j} = \delta_{a,s} \delta_{p,j}, \\ T_2 &= - \sum_{n=0}^q \sum_{i=0}^q \delta_{a,n} \delta_{p,i} \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j} \\ &= - \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j}, \\ T_3 &= \sum_{n=0}^q \sum_{i=0}^q \delta_{n,s} \delta_{i,j} \binom{q-n}{a-n} \binom{a-n}{p-i} = \binom{q-s}{a-s} \binom{a-s}{p-j}, \\ T_4 &= - \sum_{n=0}^q \sum_{i=0}^q \binom{q-n}{a-n} \binom{a-n}{p-i} \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j}. \end{aligned}$$

In the case of T_4 we note that the binomial coefficients vanish except for $s \leq n \leq a$ and $j \leq i \leq p$, when expanding them leads to the expression

$$T_4 = - \frac{(q-s)!}{(q-a)!} \sum_{n=s}^a \frac{P_{n-s+1}}{n-s+1} \sum_{i=j}^p \frac{1}{(p-i)!(a-n-p+i)!(i-j)!(n-s-i+j)!}.$$

In order to reduce this expression, will need the following identities,

$$\sum_{i=j}^p \frac{1}{(p-i)!(a-n-p+i)!(i-j)!(n-s-i+j)!} = \frac{1}{(a-n)!(n-s)!} \binom{a-s}{p-j}, \quad (\text{B.3})$$

$$\sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} = -\frac{1}{2} \delta_{b,0} - \delta_{b>0} P_{b+1}, \quad (\text{B.4})$$

and

$$\sum_{n=0}^b \frac{P_{n+1}}{(n+1)!(b-n)!} = \frac{1}{b!} - \frac{1}{2} \delta_{b,0} - \frac{P_{b+1}}{(b+1)!} \delta_{b>0}, \quad (\text{B.5})$$

where $\delta_{b>0}$ is 1 if $b > 0$ and 0 otherwise. Using these identities, which we prove at the end of this appendix,

$$\begin{aligned} T_4 &\stackrel{(\text{B.3})}{=} -\frac{(q-s)!}{(q-a)!} \binom{a-s}{p-j} \sum_{n=s}^a \frac{P_{n-s+1}}{(n-s+1)!(a-n)!} \\ &= -\frac{(q-s)!}{(q-a)!} \binom{a-s}{p-j} \sum_{n=0}^{a-s} \frac{P_{n+1}}{(n+1)!((a-s)-n)!} \\ &\stackrel{(\text{B.5})}{=} -\frac{(q-s)!}{(q-a)!} \binom{a-s}{p-j} \left[\frac{1}{(a-s)!} - \frac{1}{2} \delta_{a,s} - \frac{P_{a-s+1}}{(a-s+1)!} \delta_{a>s} \right] \\ &= -\binom{q-s}{a-s} \binom{a-s}{p-j} + \frac{1}{2} \delta_{a,s} \delta_{p,j} + \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \delta_{a>s} \end{aligned}$$

This allows us to complete the proof of Lemma 2 by proving (B.2),

$$\begin{aligned} T &= \delta_{a,s} \delta_{p,j} - \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} + \binom{q-s}{a-s} \binom{a-s}{p-j} \\ &\quad - \binom{q-s}{a-s} \binom{a-s}{p-j} + \frac{1}{2} \delta_{a,s} \delta_{p,j} + \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \delta_{a>s} \\ &= \frac{3}{2} \delta_{a,s} \delta_{p,j} - (1 - \delta_{a>s}) \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \\ &= \frac{3}{2} \delta_{a,s} \delta_{p,j} - \delta_{a \leq s} \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \\ &= \frac{3}{2} \delta_{a,s} \delta_{p,j} - \delta_{a,s} \delta_{p,j} P_1 = \delta_{a,s} \delta_{p,j}, \end{aligned}$$

in proving which we have used $P_1 = 1/2$ and the fact that SA is lower triangular (thus the only case of $a \leq s$ that we need to consider is $s = a$). This completes the proof of Lemma 2.

Proofs for the identities (B.3), (B.4) and (B.5) are given below.

Proof (B.3)

$$\begin{aligned} \sum_{i=j}^p \frac{(a-n)!(n-s)!}{(p-i)!(a-n-p+i)!(i-j)!(n-s-i+j)!} &= \sum_{i=j}^p \binom{a-n}{p-i} \binom{n-s}{i-j} \\ &= \sum_{i=0}^{p-j} \binom{a-n}{(p-j)-i} \binom{n-s}{i} = \binom{a-s}{p-j}, \end{aligned}$$

since

$$\sum_{i=0}^k \binom{n}{k-i} \binom{m}{i} = \binom{n+m}{k}.$$

□

Proof (B.4)

$$\begin{aligned} \sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} &= \sum_{n=0}^b \binom{b+1}{n} (2^{b-n+1} - 1) B_{b-n+1} \\ &= \sum_{n=0}^{b+1} \binom{b+1}{n} (2^{b-n+1} - 1) B_{b-n+1} - 0 \\ &= \left[2^{b+1} \mathcal{B}_{b+1}(1/2) - \mathcal{B}_{b+1}(1) \right] \\ &= \left[2 - 2^{b+1} - (-1)^{b+1} \right] B_{b+1}, \end{aligned} \tag{B.6}$$

where $\mathcal{B}_k(x)$ are the Bernoulli polynomials,

$$\mathcal{B}_k(x) = \sum_{i=0}^k \binom{k}{i} B_{k-i} x^i,$$

whose values at 1 and 1/2 are,

$$\mathcal{B}_k(1) = (-1)^k B_k, \quad \mathcal{B}_k(1/2) = (2^{1-k} - 1) B_k.$$

For $b = 0$, the expression (B.6) evaluates to $-1/2$, while for $b > 0$, $(-1)^{b+1}$ can be replaced by 1 since B_{b+1} vanishes for all cases when $(-1)^{b+1}$ is negative. Thus, for $b > 0$, (B.6) evaluates to $(1 - 2^{b+1}) B_{b+1}$. Using the same logic, we may multiply it by $(-1)^{b+1}$ to get $-P_{b+1}$, completing the proof. □

Proof (B.5)

$$\begin{aligned} \sum_{n=0}^b \frac{P_{n+1}}{(n+1)!(b-n)!} &= \frac{1}{(b+1)!} \sum_{n=0}^b \binom{b+1}{n+1} (-1)^{n+1} (2^{n+1} - 1) B_{n+1} \\ &= \frac{1}{(b+1)!} \left[\sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} - 2(b+1) B_1 \right] \\ &= \frac{1}{b!} + \frac{1}{(b+1)!} \sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} \\ &\stackrel{(B.4)}{=} \frac{1}{b!} - \frac{1}{2} \delta_{b,0} - \frac{P_{b+1}}{(b+1)!} \delta_{b>0}. \end{aligned}$$

Where we have used the fact that, except for the $n = 0$ case, all negative occurrences of $(-1)^{n+1}$ vanish since B_{n+1} vanishes. □

C Proof of Lemma 3—generating function

We wish to find an explicit form for the generating function

$$h(u, w, y, x) = \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \pi_{n,i}^{k,l}. \quad (3.11)$$

We start from the result of Lemma 2, substituting (3.9) in (3.11) and splitting $\pi_{n,i}^{k,l}$ into eight parts for convenience,

$$\pi_{n,i}^{k,l} = \frac{1}{2} \sum_{j=1}^8 p_j(k, l, n, i), \quad (C.1)$$

where

$$\begin{aligned} p_1(k, l, n, i) &= \delta_{i,0} \binom{k}{n}, \\ p_2(k, l, n, i) &= \delta_{i,0} \binom{k+l}{n}, \\ p_3(k, l, n, i) &= \binom{k}{i} \binom{k-i}{n-i}, \\ p_4(k, l, n, i) &= \binom{k}{i} \binom{k+l-i}{n-i}, \\ p_5(k, l, n, i) &= - \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k}{n-r}, \\ p_6(k, l, n, i) &= - \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k+l}{n-r}, \\ p_7(k, l, n, i) &= - \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i-j} \binom{k}{j} \binom{k-j}{n-r-j}, \\ p_8(k, l, n, i) &= - \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i-j} \binom{k}{j} \binom{k+l-j}{n-r-j}, \end{aligned}$$

are obtained after a change of variables, $r = n - s$, and noting the facts that $\binom{0}{i-j} = \delta_{i,j}$ and that $\binom{n}{k}$ vanishes for $k < 0$ when $n \geq 0$. We will simplify the corresponding parts of $h(u, w, y, x)$,

$$h_j(u, w, y, x) = \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i p_j(k, l, n, i),$$

combining them to find an expression for the generating function $h = \sum_{j=1}^8 h_j/2$. In this pursuit, we will repeatedly use a few results,

$$\sum_{r=0}^{\infty} \frac{B_r}{r!} x^r = \frac{x}{e^x - 1}, \quad (C.2)$$

which is well known,

$$\begin{aligned}
 \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} x^r &= \frac{1}{x} \sum_{r=0}^{\infty} \frac{P_r}{r!} x^r = \frac{1}{x} \sum_{r=0}^{\infty} \frac{(-1)^r (2^r - 1) B_r}{r!} x^r \\
 &= \frac{1}{x} \sum_{r=0}^{\infty} \frac{B_r}{r!} (-2x)^r - \frac{1}{x} \sum_{r=0}^{\infty} \frac{B_r}{r!} (-x)^r \\
 &= - \left(\frac{2}{e^{-2x} + 1} - \frac{1}{e^{-x} + 1} \right) = \frac{e^x}{e^x + 1},
 \end{aligned} \tag{C.3}$$

where we use the fact that $P_0 = 0$, and

$$\begin{aligned}
 \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+r}{l} u^l w^{k+n-l} &= \sum_{k=0}^{\infty} w^{k+n} \sum_{l=0}^{\infty} \binom{k+r}{l} \left(\frac{u}{w} \right)^l = \sum_{k=0}^{\infty} w^{k+n} \left(1 + \frac{u}{w} \right)^k + r \\
 &= w^{n-r} (w+u)^r \sum_{k=0}^{\infty} (w+u)^k = \frac{w^{n-r} (w+u)^r}{1 - (w+u)}.
 \end{aligned} \tag{C.4}$$

Two standard tricks for exchanging summations that we exploit are

$$\sum_{n=0}^{\infty} \sum_{r=0}^n \alpha_{n,r} = \sum_{r=0}^{\infty} \sum_{n=0}^{\infty} \alpha_{n+r,r}, \tag{C.5}$$

$$\sum_{k=0}^{\infty} \sum_{n=0}^{k+l} \alpha_{n,k,l} = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \alpha_{n,k+n-l,l}. \tag{C.6}$$

With these tools, we proceed to seek expressions for the generating functions h_j , writing h_j as shorthand for $h_j(u, w, y, x)$.

$$\begin{aligned}
 h_1 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \delta_{i,0} \binom{k}{n} \\
 &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \binom{k+l-n}{l} \frac{y^n}{n!} = \sum_{n=0}^{\infty} \frac{y^n}{n!} \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k}{l} u^l w^{k+n-l} \right] \\
 &= \frac{1}{1 - (w+u)} \sum_{n=0}^{\infty} \frac{(yw)^n}{n!} = \frac{e^{yw}}{1 - (w+u)},
 \end{aligned}$$

$$\begin{aligned}
 h_2 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \delta_{i,0} \binom{k+l}{n} \\
 &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \binom{k+l}{l} \frac{y^n}{n!} = \sum_{n=0}^{\infty} \frac{y^n}{n!} \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n}{l} u^l w^{k+n-l} \right] \\
 &= \frac{1}{1 - (w+u)} \sum_{n=0}^{\infty} \frac{(y(w+u))^n}{n!} = \frac{e^{y(w+u)}}{1 - (w+u)},
 \end{aligned}$$

$$\begin{aligned}
h_3 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \binom{k}{i} \binom{k-i}{n-i} \\
&= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \frac{y^n}{n!} \binom{k+l-n}{l} \left[\sum_{i=0}^n \binom{n}{i} x^i \right] \\
&= \sum_{n=0}^{\infty} \frac{(y(1+x))^n}{n!} \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k}{l} u^l w^{k+n-l} \right] \\
&= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{(yw(1+x))^n}{n!} = \frac{e^{yw(1+x)}}{1-(w+u)},
\end{aligned}$$

$$\begin{aligned}
h_4 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \binom{k}{i} \binom{k+l-i}{n-i} \\
&= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \frac{y^n}{n!} \sum_{i=0}^n \binom{n}{i} x^i \binom{k+l-i}{l} \\
&= \sum_{n=0}^{\infty} \frac{y^n}{n!} \sum_{i=0}^n \binom{n}{i} x^i \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n-i}{l} u^l w^{k+n-l} \right] \\
&= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{y^n}{n!} \sum_{i=0}^n \binom{n}{i} x^i w^i (w+u)^{n-i} \\
&= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{(y((w+u)+xw))^n}{n!} = \frac{e^{y(w+u)+xyw}}{1-(w+u)},
\end{aligned}$$

$$\begin{aligned}
h_5 &= - \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k}{n-r} \\
&= - \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \sum_{i=0}^r x^i \frac{P_{r+1}}{(r+1)!(n-r)!} \binom{k+l-n+r}{l} \binom{r}{i} \\
&= - \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+r}{l} u^l w^{k+n-l} \right] \left[\sum_{i=0}^r \binom{r}{i} x^i \right],
\end{aligned}$$

where we have changed summation limits on n using (C.6), and for i since $r \leq n$ and $\binom{r}{i} = 0$ for $i > r \geq 0$. Using (C.4) and changing limits again using (C.5),

$$\begin{aligned}
h_5 &= - \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} w^{n-r} ((w+u)(1+x))^r \\
&= - \frac{1}{1-(w+u)} \left[\sum_{n=0}^{\infty} \frac{(yw)^n}{n!} \right] \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} (y(w+u)(1+x))^r \\
&= - \frac{e^{yw}}{1-(w+u)} \frac{e^z}{e^z + 1},
\end{aligned}$$

where we have defined $z = y(w + u)(1 + x)$ for convenience.

$$\begin{aligned}
 h_6 &= - \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k+l}{n-r} \\
 &= - \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} \binom{k+l}{l} \left[\sum_{i=0}^r \binom{r}{i} x^i \right] \\
 &= - \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} (1+x)^r \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n}{l} u^l w^{k+n-l} \right] \\
 &= - \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} (y(w+u))^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} (1+x)^r \\
 &= - \frac{1}{1-(w+u)} \left[\sum_{n=0}^{\infty} \frac{(y(w+u))^n}{n!} \right] \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} (y(w+u)(1+x))^r \\
 &= - \frac{e^{y(w+u)}}{1-(w+u)} \frac{e^z}{e^z + 1}.
 \end{aligned}$$

After manipulating binomial coefficients, we change summation limits (C.6) and use (C.4),

$$\begin{aligned}
 h_7 &= - \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} \binom{k+l-n+r}{l} \left[\sum_{i=0}^n \binom{r}{i-j} x^i \right] \\
 &= - \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+r}{l} u^l w^{k+n-l} \right] \left[\sum_{i=j}^{r+j} \binom{r}{i-j} x^i \right] \\
 &= - \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} w^{n-r} (w+u)^r x^j (1+x)^r,
 \end{aligned}$$

since $\binom{r}{i-j}$ vanishes unless $j \leq i \leq r+j$ and $r+j \leq n$. Exchanging limits twice using (C.5), once for r and once for j ,

$$\begin{aligned}
 h_7 &= - \frac{1}{1-(w+u)} \sum_{r=0}^{\infty} \sum_{n=0}^{\infty} (yw)^n \sum_{j=0}^n \frac{P_{r+1}}{(r+1)!j!(n-j)!} z^r x^j \\
 &= - \frac{1}{1-(w+u)} \left[\sum_{n=0}^{\infty} \frac{(yw)^n}{n!} \right] \left[\sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} z^r \right] \left[\sum_{j=0}^{\infty} \frac{(ywx)^j}{j!} \right] \\
 &= - \frac{e^{yw(1+x)}}{1-(w+u)} \frac{e^z}{e^z + 1}.
 \end{aligned}$$

Along similar lines we manipulate

$$\begin{aligned}
h_8 &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{-P_{r+1}}{(r+1)!j!(n-r-j)!} \binom{k+l-j}{l} \left[\sum_{i=0}^n \binom{r}{i-j} x^i \right] \\
&= \sum_{n=0}^{\infty} \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{-P_{r+1} y^n}{(r+1)!j!(n-r-j)!} \left[\sum_{i=j}^{r+j} \binom{r}{i-j} x^i \right] \left[\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n-j}{l} u^l w^{k+n-l} \right] \\
&= -\frac{1}{1-(w+u)} \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} x^j (1+x)^r w^j (w+u)^{n-j},
\end{aligned}$$

since $\binom{r}{i-j}$ vanishes unless $j \leq i \leq r+j$ and $r+j \leq n$. Once again, exchanging limits twice using (C.5), once for r and once for j ,

$$\begin{aligned}
h_8 &= -\frac{1}{1-(w+u)} \sum_{r=0}^{\infty} \sum_{n=0}^{\infty} y^{n+r} \sum_{j=0}^n \frac{P_{r+1}}{(r+1)!j!(n-j)!} x^j (1+x)^r w^j (w+u)^{n+r-j} \\
&= -\frac{1}{1-(w+u)} \left[\sum_{n=0}^{\infty} \frac{(y(w+u))^n}{n!} \right] \left[\sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} z^r \right] \left[\sum_{j=0}^{\infty} \frac{(xyw)^j}{j!} \right] \\
&= -\frac{e^{y(w+u)+xyw}}{1-(w+u)} \frac{e^z}{e^z+1}.
\end{aligned}$$

These expressions are then combined to form the full generating function

$$\begin{aligned}
2h(u, w, y, x) &= \frac{e^{yw}}{1-(w+u)} + \frac{e^{y(w+u)}}{1-(w+u)} + \frac{e^{yw(1+x)}}{1-(w+u)} + \frac{e^{y(w+u)+xyw}}{1-(w+u)} \\
&\quad - \frac{e^{yw}}{1-(w+u)} \frac{e^z}{e^z+1} - \frac{e^{y(w+u)}}{1-(w+u)} \frac{e^z}{e^z+1} \\
&\quad - \frac{e^{yw(1+x)}}{1-(w+u)} \frac{e^z}{e^z+1} - \frac{e^{y(w+u)+xyw}}{1-(w+u)} \frac{e^z}{e^z+1} \\
&= \frac{e^{yw} + e^{y(w+u)} + e^{yw(1+x)} + e^{y(w+u)+xyw}}{1-(w+u)} \left[1 - \frac{e^z}{e^z+1} \right] \\
&= \frac{1}{1-(w+u)} \left[\frac{e^{yw} + e^{y(w+u)} + e^{yw(1+x)} + e^{y(w+u)+xyw}}{e^z+1} \right].
\end{aligned}$$

For further simplification, consider $g = 2h(u, w, 2y, x)(1 - (w + u))$, letting $a = uy, b = wy, c = uxy$ and $d = wxy$,

$$\begin{aligned}
 g &= \frac{e^{2yw} + e^{2y(w+u)} + e^{2yw(1+x)} + e^{2y(w+u)+2xyw}}{e^{2yw+2ywx+2yu+2yux} + 1} \\
 &= \frac{e^{2b} + e^{2a+2b} + e^{2b+2d} + e^{2a+2b+2d}}{e^{2a+2b+2d+2c} + 1} \\
 &= \frac{e^{-a+b-c-d} + e^{a+b-c-d} + e^{-a+b-c+d} + e^{a+b-c+d}}{e^{a+b+c+d} + e^{-a-b-c-d}} \\
 &= e^{b-c} \frac{e^{-a-d} + e^{a-d} + e^{-a+d} + e^{a+d}}{e^{a+b+c+d} + e^{-a-b-c-d}} \\
 &= e^{b-c} \frac{(e^{-a} + e^a)(e^{-d} + e^d)}{e^{a+b+c+d} + e^{-a-b-c-d}} \\
 &= 2e^{b-c} \frac{\cosh(a) \cosh(d)}{\cosh(a + b + c + d)}.
 \end{aligned}$$

This brings us to the desired form of the generating function,

$$h(u, w, y, x) = \frac{\exp((wy - uxy)/2)}{1 - (w + u)} \frac{\cosh(uy/2) \cosh(wxy/2)}{\cosh(y(u + w)(1 + x)/2)}, \quad (3.12)$$

and completes the proof of Lemma 3.